



HAL
open science

A propos du MIMO massif dans un contexte de sobriété numérique

Philippe Ciblat

► **To cite this version:**

Philippe Ciblat. A propos du MIMO massif dans un contexte de sobriété numérique. GRETSI, Sep 2022, Nancy, France. hal-03700928

HAL Id: hal-03700928

<https://hal.science/hal-03700928v1>

Submitted on 21 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A propos du MIMO massif dans un contexte de sobriété numérique

Philippe CIBLAT

LTCI, Télécom Paris, Institut Polytechnique de Paris
19 place Marguerite Perey, 91120 Palaiseau, France
ciblat@telecom-paris.fr

Résumé – Nous étudions l’apport possible de la technique du MIMO massif d’un point de vue énergétique lors de l’introduction d’une nouvelle génération de téléphonie mobile. Nous montrons que l’intérêt est flagrant seulement si les énergies durant la phase d’usage sont considérées. En revanche, en prenant en compte uniquement la phase de fabrication, l’intérêt dépend fortement de l’évolution de trafic envisagé (frugalité ou expansion) et aussi de la durée de renouvellement d’une génération. A titre d’exemple, avec une croissance de trafic de 10% par an, le renouvellement doit intervenir au bout de 12 ans. Sans croissance de trafic, l’impact de l’installation d’une génération plus efficace énergétiquement durant la phase d’usage et donc à base de MIMO massif est négatif.

Abstract – We analyze the pros and cons of the massive MIMO with an energetic point of view, when advocated and launched in a next wireless system generation. Massive MIMO is of great interest when only the energy during the use phase is considered. In contrast, by taking into account the energy consumed during the manufacturing phase, the benefit for changing the physical layer (for instance, by introducing the massive MIMO) is questionable and strongly depends on the traffic model and the updating duration. According to our model, with an increase of 10% a year, the duration of a generation should be at least 12 years. Without traffic increase, manufacturing and carrying out a new more energy-efficient (as massive MIMO) generation is useless.

1 Introduction

A chaque renouvellement de génération d’un système de téléphonie mobile, une nouvelle couche physique est mise en place (CDMA pour la 3G, OFDM pour la 4G et MIMO massif pour la 5G). Cette nouvelle couche (en plus de l’augmentation inhérente de la largeur de bande) permet d’augmenter le débit et d’améliorer l’efficacité énergétique de la phase d’usage, c’est-à-dire, le nombre de bits transportés par Joule (ce constat est juste pour le passage de la 4G vers la 5G mais dépend néanmoins de la manière d’apprécier la phase d’usage ; nous reviendrons sur ce point plus tard). Une dernière avancée entre la 4G et la 5G est protocolaire et correspond à la mise en veille des stations de base en l’absence de trafic.

De nombreux travaux se sont intéressés à l’optimisation de l’efficacité énergétique en phase d’usage. Néanmoins avec ce type de critère, on ne maîtrise pas le point de fonctionnement qui peut induire une sur-qualité en suggérant un débit très élevé et finalement une possibilité d’effet rebond. C’est pourquoi dans ce papier, on reprend un point de vue plus classique en examinant la consommation énergétique pour un point de fonctionnement donné, et notamment un point en lien avec la sobriété énergétique, c’est-à-dire, pour un débit raisonnable par utilisateur et un nombre raisonnable d’utilisateurs. Ce contexte de sobriété conduit à prôner un réseau conçu juste pour satisfaire des besoins essentiels à définir collectivement et démocratique-

ment (en anglais, « *just enough network* » ; en français « *réseau juste suffisant* ») et donc à rebours de l’hybris actuelle que l’on constate dans les présentations sur la future 6G.

Avec ce regard différent, nous analysons l’intérêt du MIMO massif en prenant en compte d’abord les différents types d’énergies consommées durant la phase d’usage (puissance de transmission, puissance des circuits analogiques, puissance de traitement numérique, etc) puis en y ajoutant l’énergie consommée durant la phase de fabrication ce qui nous permettra d’étudier l’impact de la durée d’amortissement.

L’article est organisé ainsi : dans la Section 2, on revisite le critère d’efficacité énergétique à l’aune du concept de réseau juste suffisant. A la Section 3, nous présentons les différentes énergies consommées par un MIMO massif. Le nombre optimal d’antennes pour un trafic donné est caractérisé en Section 4. Les résultats numériques sont discutés en Section 5.

2 Réflexion sur l’efficacité énergétique

Dans cette section, nous livrons quelques éléments de réflexion critique sur la validité du critère de l’efficacité énergétique, en intégrant plusieurs dimensions : notamment mise en veille ou pas du réseau et effet rebond.

Dans le cadre d’une communication point-à-point avec une seule antenne, le débit R (en bits/s) admet la forme suivante, $R = B \log_2(1 + P/(BN_0))$ avec P la puissance de transmission, B la largeur de bande et N_0 le niveau du bruit.

Nous considérons dans la suite deux manières de dimension-

Ce travail n’est soutenu par aucun contrat. Je tiens à remercier Sophie Benkő, Marceau Coupechoux, Chadi Jabbour, Germain Pham et Christine Tavernier pour leur discussion et commentaires fructueux.

ner le réseau :

- la puissance consommée se décompose en deux éléments : le premier est la puissance de transmission P et la puissance de circuiterie P_{circuit} (dont on détaillera les composantes en Section 3). Par conséquent, l'efficacité énergétique (EE) s'écrit

$$EE = \frac{B \log_2 \left(1 + \frac{P}{BN_0} \right)}{P + P_{\text{circuit}}}$$

ce qui conduit en la maximisant au point de fonctionnement suivant en terme de débit [1]

$$R_{EE}^* = \frac{B}{\ln(2)} \left(W \left(\frac{P_{\text{circuit}}}{eBN_0} - \frac{1}{e} \right) + 1 \right)$$

et de puissance de transmission

$$P_{EE}^* = BN_0 \left(e^{W \left(\frac{P_{\text{circuit}}}{eBN_0} - \frac{1}{e} \right) + 1} - 1 \right),$$

où $W(\cdot)$ est la fonction W de Lambert.

- on adapte la puissance de transmission pour émettre au débit R_E^* . Ainsi, on a

$$P_E^* = BN_0 (2^{R_E^*/B} - 1).$$

Avant d'aller plus loin, on peut remarquer que si P_{circuit} tend vers zéro (resp. $+\infty$), alors le débit R_{EE}^* tend également vers zéro (resp. $+\infty$). Plus la circuiterie des appareils est optimisée et moins le système aura un débit élevé et donc cela conduit à une grande sobriété numérique !

Revenons maintenant au cœur du sujet : on souhaite transmettre un fichier de taille F_x avec $x \in \{EE, E\}$ car on s'autorise des fichiers de taille différente suivant l'optimisation choisie. Pour simplifier, on a choisi P_{circuit} tel que $R_{EE}^* > R_E^*$. L'énergie consommée pour transmettre ce fichier s'écrit

- dans le cas d'un débit pré-défini par

$$E_E^* = \frac{F_E BN_0 (2^{R_E/B} - 1)}{R_E^*} + \frac{F_E P_{\text{circuit}}}{R_E^*}$$

où F_E/R_E^* correspond à la durée de la transmission.

- et dans le cas d'un débit obtenu avec l'optimisation de l'EE par

$$E_{EE}^* = \frac{F_{EE} BN_0 \left(e^{W \left(\frac{P_{\text{circuit}}}{eBN_0} - \frac{1}{e} \right) + 1} - 1 \right)}{R_{EE}^*} + \frac{F_{EE} P_{\text{circuit}}}{R_{EE}^*} + \nu F_{EE} P_{\text{circuit}} \left(\frac{1}{R_E^*} - \frac{1}{R_{EE}^*} \right)$$

où $\mu := 1 - \nu$ où ν est un facteur de veille. Si $\nu = 1$, le système est en veille dès que l'unique fichier considéré est transmis. Si $\nu = 0$, il n'y a pas de veille ou bien le système a un autre fichier à transmettre (dans le cas d'un modèle continu de trafic). Si $\nu \in]0, 1[$, un système de veille partielle est mis en œuvre.

Dans les deux cas, le premier terme de la partie droite de l'équation désigne l'énergie dédiée à la transmission.

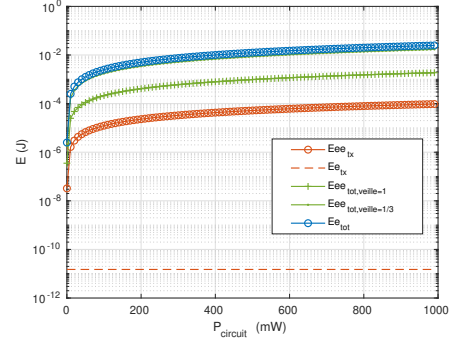


FIGURE 1 – Energie transmise et totale en fonction de P_{circuit} .

Sur la figure 1, on trace l'énergie transmise et totale (avec $\nu = 0$, $\nu = 1$ et $\nu = 1/3$) en fonction de P_{circuit} . Les simulations ont été effectuées avec $N_0 = -140$ dBm/Hz, $B = 20$ MHz, $R_E^* = 40$ Mbits/s et $F_E = F_{EE} = 1$ Mbits. On remarque que si la mise en veille est parfaite, le critère de l'efficacité énergétique est extrêmement profitable. En revanche, sans mise en veille, minimiser la puissance de transmission émise devrait primordial et alors adapter le débit à celui réellement nécessaire pour la qualité de service est une meilleure option. En effet il y a un compromis entre le temps sur lequel on considère la communication (le critère d'efficacité énergétique a tendance à induire un débit élevé et donc un temps de transmission court ce qui à puissance constante conduit à baisser l'énergie consommée) et l'énergie consommée pour transmettre (petite à débit faible).

Dans un contexte de mise en veille parfaite ($\nu = 1$), transmettre au débit imposé par le critère d'efficacité énergétique est avantageux par rapport au débit minimal imposé par la qualité de service. Par effet rebond, il est courant qu'une meilleure performance induise une augmentation du trafic notamment par l'augmentation de la taille des fichiers (meilleure qualité d'image par exemple). Par conséquent, on souhaite connaître l'augmentation admissible pour le fichier de taille F_{EE} émis par le débit R_{EE}^* par rapport au fichier de taille F_E émis par le débit R_E^* . La simulation a été effectuée avec les mêmes paramètres que ceux de la figure 1 et $P_{\text{circuit}} = 1$ W. On a remarqué qu'une hausse d'un facteur 10 de la taille du fichier obère le gain obtenu par le critère d'efficacité énergétique et donc une politique de sobriété s'avère nécessaire.

3 Modèle réaliste du système

Dans cette section, on va analyser l'intérêt du MIMO massif (avec N antennes) d'un point de vue énergétique dans le contexte d'un réseau juste suffisant. C'est pourquoi, au lieu de raisonner sur l'efficacité énergétique, on va considérer l'énergie minimale à consommer pour écouler un certain trafic caractérisé par le nombre d'utilisateurs K et le débit requis R .

Avant d'aller plus loin, décrivons le système : on considère une liaison descendante mono-cellulaire entre une station de

base à antennes multiples et des utilisateurs à antenne unique. A chaque trame, au maximum $(N - 1)$ utilisateurs peuvent être servis. Par conséquent les utilisateurs doivent être répartis sur $S := \lceil K/(N - 1) \rceil$ trames dans lesquelles il y a $K_u := \lceil K/S \rceil$ utilisateurs actifs maximum. En conséquence, un utilisateur lorsqu'il est actif requiert un débit $R_u := SR$ puisqu'il n'utilise qu'une trame sur S . Notons que le MIMO massif joue finalement aussi le rôle avantageux de technique d'accès multiple non-orthogonal permettant de satisfaire plus d'utilisateurs dans une même trame avec un débit amoindri pour chacun. On suppose que le trafic est continu (finalement $\nu = 1$ dans la Section précédente). Ainsi si un utilisateur est amené à disparaître, il est remplacé par un autre localisé au même endroit.

Les utilisateurs sont localisés de manière uniforme dans une couronne de petit rayon r_{\min} et de grand rayon r_{\max} . Le modèle d'atténuation est le suivant : $x \mapsto d_0/|x|^\kappa$ avec x la distance au centre de la couronne, d_0 la constante d'atténuation et κ le facteur d'atténuation.

Nous listons maintenant les différentes énergies impliquées : comme le système est toujours actif, on peut raisonner en puissance. Ainsi, la puissance totale consommée pour une liaison descendante se décompose de la manière suivante :

$$P_{\text{tot}} = P_{\text{tx}} + P_{\text{fab}} + P_{\text{circuit}} \quad (1)$$

où

- P_{tx} est la puissance de transmission qui vaut si un précodeur à Forçage à Zéro a été mis en place [1]

$$P_{\text{tx}} = \frac{BN_0}{\eta} \frac{2^{R_u/B} - 1}{N - K_u} \mathcal{D}K_u$$

avec η l'efficacité de l'amplificateur de puissance de la station de base et $\mathcal{D} := \mathbb{E}[|x|^\kappa]/d_0$ une constante dépendant du modèle de propagation.

- P_{fab} est la puissance nécessaire à la fabrication des appareils (station de base –bts– et téléphones mobiles –ue–). Ainsi $P_{\text{fab}} = NP_{\text{fab,bts}} + KP_{\text{fab,ue}}$ où $P_{\text{fab,bts}}$ est la puissance nécessaire pour construire une station de base à antenne multiple. Généralement on a connaissance de l'énergie consommée $E_{\text{fab,d}}$ ($d \in \{\text{bts, ue}\}$) nécessaire pour fabriquer un appareil avec une antenne. On utilise alors la règle suivante : $P_{\text{fab,d}} = E_{\text{fab,d}}/(365 \times 24 \times 3600 \times A)$ avec A le nombre d'années d'amortissement considéré.
- P_{circuit} est la puissance de circuiterie de la station de base et des téléphones. On a

$$P_{\text{circuit}} = P_{\text{fix}} + P_{\text{tc}} + P_{\text{ce}} + P_{\text{cd}} + P_{\text{bh}} + P_{\text{lp}}$$

avec

- P_{fix} la puissance consommée en veille par la station de base (typiquement le système de refroidissement).
- P_{tc} la puissance des chaînes d'émetteur/récepteur qui vaut $P_{\text{tc}} = NP_{\text{bts}} + P_{\text{lo}} + KP_{\text{ue}}$ avec l'unique oscillateur local P_{lo} à la station de base et P_{bts} et P_{ue} respectivement la consommation des circuits (convertisseurs, filtres, mixeurs) à une antenne de la station de base et au téléphone. Comme tous les utilisateurs restent allumés à chaque trame, P_{ue} est multipliée par K et non par K_u .

- P_{ce} la puissance pour estimer les canaux. On a

$$P_{\text{ce}} = \frac{B}{U} \frac{4\tau K_u^2}{L_{\text{ue}}}$$

où U le nombre de symboles sur lequel le canal est constant, τ le pourcentage de la trame dédiée à la séquence d'apprentissage et L_{ue} le nombre d'opérations possibles par Watt (mesuré en flop/W).

- P_{cd} la puissance de codage et décodage. Quelques articles seulement modélisent cette puissance sans converger vers un consensus. Selon [1], il y a une proportionnalité avec le débit (alors que d'autres articles suggèrent une décroissance avec une certaine distance à la capacité de Shannon). Ainsi, $P_{\text{cd}} = K_u R_u (P_c + P_d)$ avec P_c et P_d les puissances unitaires de codage et décodage respectivement.
- P_{bh} la puissance du réseau-cœur dépendant du trafic. Si P_{bt} est la puissance unitaire, alors

$$P_{\text{bh}} = K_u R_u P_{\text{bt}}.$$

- P_{lp} la puissance liée au traitement du signal (ex., les précodeurs linéaires pour la formation de voie). Alors $P_{\text{lp}} = P_{\text{mc}} + P_{\text{mo}}$ avec P_{mc} la puissance nécessaire à l'application des codeurs linéaires qui vaut

$$P_{\text{mc}} = B \left(1 - \frac{\tau K_u}{U} \right) \frac{2NK_u}{L_{\text{bts}}},$$

et P_{mo} la puissance nécessaire à l'optimisation de ces codeurs linéaires (en supposant qu'un Forçage à Zéro a été mis en œuvre) qui vaut

$$P_{\text{mc}} = \frac{B}{U} \left(\frac{K_u^3}{3L_{\text{bts}}} + \frac{3NK_u^2 + NK_u}{L_{\text{bts}}} \right)$$

avec L_{bts} le nombre d'opérations possibles par Watt.

4 Optimisation du système

Pour analyser l'intérêt de MIMO massif, on souhaite résoudre le problème d'optimisation suivant

$$N^* = \arg \min_N P_{\text{tot}}(N, R, K)$$

où $P_{\text{tot}}(N, R, K)$ est donnée par l'équation (1). Ce problème se résout simplement par une recherche mono-dimensionnelle sur la variable N .

Avant d'aller plus loin, on souhaite connaître les ordres de grandeur des différents termes impliqués dans l'équation (1) afin de la simplifier. Le tableau 1 regroupe les différentes valeurs des constantes introduites jusqu'à présent dans le papier. La plupart provient de [1].

D'après [2], la fabrication et installation d'une station de base 4G avec 6 antennes libère 10 tCO₂e. De plus, le mix énergétique français et scandinave donne environ 100 gCO₂e/kWh. Par une simple règle de trois, on obtient la valeur indiquée pour $E_{\text{fab,bts}}$. Pour la fabrication d'un téléphone mobile, on utilise la valeur donnée par [3].

B	20 MHz	$E_{\text{fab,bts}}$	60 GJ
N_0	-140 dBm/Hz	$E_{\text{fab,ue}}$	0,175 GJ
η	0,39	P_{fix}	18 W
U	1800	P_{bts}	1 W
τ	1	P_{lo}	2 W
L_{ue}	5 Gflops/W	P_{ue}	0,1 W
L_{bts}	12,8 Gflops/W	P_{c}	0,1 W/(Gbits/s)
d_0	$10^{-3,53}$	P_{d}	0,8 W/(Gbits/s)
κ	3,76	P_{bt}	0,25 W/(Gbits/s)

TABLE 1 – Valeurs considérées.

Dans le tableau 2, on donne les ordres de grandeur des puissances en Watt (sauf celles de la fabrication) pour deux configurations : la première est avec $K = 5$ et $R = 10$ kbits/s (qui a donné $N^* = 20$) et la seconde avec $K = 50$ et $R = 10$ Mbits/s (qui a donné $N^* = 1486$).

P_{tx}	P_{fix}	P_{tc}	P_{ce}	P_{cd}	P_{bh}	P_{lp}
13	18	22,5	2.10^{-4}	5.10^{-5}	15.10^{-6}	0,3
1660	18	1491	0.02	0,5	0,15	235

TABLE 2 – Valeurs des différentes puissances (en W).

En ne conservant que les termes dominants du tableau 2, (c'est-à-dire, P_{tx} , P_{tc} , P_{lp}), on montre que la puissance totale a pour allure (à K fixé et $N > K$), $P = \alpha_K/(N - K) + \beta_K N$ avec $\alpha_K = aK$ et $\beta_K := b_0 + b_1 K + b_2 K^2$ où a , b_0 , b_1 et b_2 sont des constantes. Et a et b_2 sont positifs. Par conséquent,

$$N^* = K + \sqrt{\beta_K/\alpha_K}$$

est une fonction croissante en K . Ainsi plus le nombre d'utilisateurs est dense, et plus le MIMO massif est intéressant malgré sa puissance de circuiterie qui augmente. Ce type de résultat est similaire à [1] obtenu avec le critère d'efficacité énergétique.

5 Résultats numériques

Sur la figure 2, on trace N^* en fonction de R pour différentes valeurs de K . On n'y considère pas l'énergie induite par la fabrication. Le nombre d'antennes croît avec l'exigence en débit et donc une nouvelle génération qui serait conçue pour absorber ou créer des très haut débits nécessite un recours au MIMO massif. On remarque qu'on passe dans l'intervalle de R considéré de 20 antennes à plus de 1000.

Sur la figure 3, on trace E_{tot} en fonction de la durée d'usage de la technique utilisée (ci-dessous appelée durée d'amortissement). L'énergie induite par la fabrication est maintenant prise en compte. On appelle trafic le terme $T = KR$. S'il augmente d'un facteur $(1 + t/100)$ par an où t est le taux de croissance en pourcentage, il est équiréparti entre le nombre d'utilisateurs (qui augmente donc de $\sqrt{1 + t/100}$) et le débit (qui augmente du même facteur). La technique dite *equiv-4G* utilise 4 antennes alors que celle dite *equiv-5G* en a 100. De plus

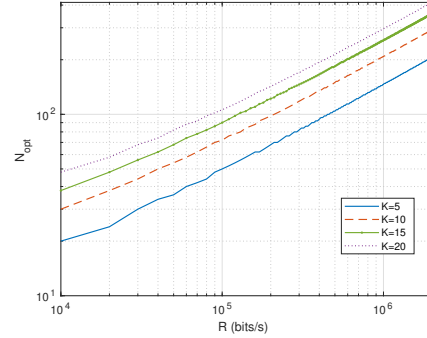


FIGURE 2 – N^* en fonction de R .

à l'instant 1 de la figure, l'*equiv-5G* est mise en place alors que l'*equiv-4G* a déjà 10 ans de fonctionnement. On remarque que sans hausse de trafic, l'*equiv-5G* n'a pas intérêt à être installée et que la durée actuelle d'amortissement qui tourne autour de 10 ans est obtenue avec une hausse d'un peu moins de 10%. En revanche avec les hausses enregistrées actuelles, l'*equiv-5G* est intéressante au bout de 5 ans. Nos valeurs de fabrication induisent une puissance de plusieurs centaines de Watt par station de base avec un amortissement de 10 ans à comparer avec le tableau 2 et ainsi la fabrication est comprise entre les valeurs de P_{tx} et donc globalement comparable alors que d'autres données mettent plus en avant la fabrication jusqu'à une hauteur du 90% et donc désavantagent l'*equiv-5G*.

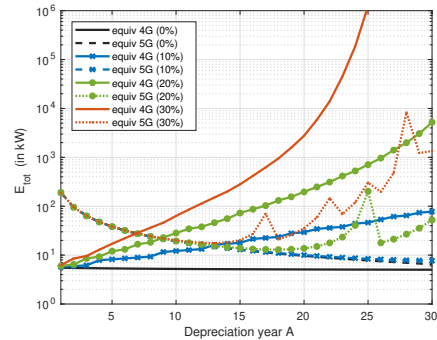


FIGURE 3 – Energie en fonction de la durée d'amortissement.

Références

- [1] E. Björnson, L. Sanguinetti, J. Hoydis et M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems : is massive MIMO the answer?", *IEEE Trans. on Wireless Com.*, vol. 14, no. 6, pp. 3059-3024, Juin 2015.
- [2] P. Bergmark, "Lice Cycle Assessment of an LTE base station based on primary data", *ETSI Workshop on ICT Energy Efficiency and Environmental Sustainability*, Sophia-Antipolis (France), Juin 2015.
- [3] J. Yu, E. Williams et M. Ju "Analysis of material and energy consumption of mobile phone in China", *Energy Policy*, no. 38, pp. 4135 - 4141, 2010.